

Rakib Khan
Rk704
Supervisor: Biagio Forte

Space Weather and Machine Learning Log Book

Project Title: Machine Learning to Identify Different Space Weather Conditions – (ID 218)

Name: Rakib Khan

Student number (candidate number): 11227

BEng interim individual project report

Academic year: 3

Supervisor name: Biagio Forte

11/02/2021 (Week 1) - Starting up and Setting up

Searched for and found research articles and literature on previous research and projects relating to applying Machine Learning to the field of Space Weather.

Emailed and organised meeting with supervisor.

Met with supervisor and discussed the project, the feasibility of the project and other aspects.

Set up a development environment using python (Jupyter Notebooks).

Familiarised myself with Jupyter Notebooks (watched a few tutorials and tested it out)

19/02/2021 (Week 2) - The Search for Data

Started searching for data, found a site called EarthData (by NASA).

Seems to be a lot of data that seemed promising, but most of it in .nc format.

Tried converting .nc to csv for a while, didn't work, time and effort wasted.

Met with supervisor and stop trying to change .nc to .csv and instead use data in its .nc format.

Downloaded some more data.

Turns out that all this data was in IONEX format.

Double checked with supervisor and found out this wasn't the data I should be focusing on.

I should've been focusing on radio occultation data from COSMIC.

Tried searching again and again for this data on NASA's site, google, github, but to no avail.

My current struggles are with finding the data.

Rakib Khan
Rk704
Supervisor: Biagio Forte

Finally found the data on UCAR COSMIC, thanks to help from my supervisor.

26/02/2021 (Week 3) - Netcdf Files

Newly acquired data are netcdf files (.nc files) and here restarts my problem of trying to open .nc files in python.

The netcdf4 module keeps throwing an error in jupyter notebook.

I deleted and reinstalled it, still throws the same error.

Takes forever waiting for anaconda navigator to update the netcdf4 module, hasn't finished once.

Uninstalled the entire program.

Removed anaconda, jupyter and everything from my laptop and reinstalled.

Was a very time consuming heavy process.

But now I can't even find netcdf4 when searching on anaconda.

I try and install netcdf4 using the anaconda prompt.

It finally installed successfully.

05/03/2021 (Week 4) - Comprehending the Data

Talked to supervisor, need to do more background reading to properly understand the data in the files.

Did more reading on literature.

Took another look at the data files, the variables now seem to make more sense.

Tried logging out the variables such as TEC, time etc.

Played around and understood the variables in more depth by checking the data type, size of array etc.

12/03/2021 (Week 5) - Plotting with Basemap

Now need to plot and visualise the data.

Looking through python packages and libraries for this task. Most popular and easy one seems to Basemap with many tutorials available.

Trying to install Basemap seems to be another cumbersome issue similar to the previous problem of trying to install the netcdf4 package.

Trying multiple different ways of installing basemap but to no success...

Out of pure luck, I stumble across a random stackoverflow post which suggests to insert this one line of code before my import line code:

```
os.environ["PROJ_LIB"] = "C:\\Users\\rakib\\anaconda3\\Library\\share"
```

and for some reason, this actually works, I have no idea why and will probably never know.

19/03/2021 (Week 6) – Debugging Plotting

Watching and reading tutorials and documentation on basemap.

Got some basic plots going, things seem to be starting to come together.

Basemap starts throwing a random import error despite the fix from earlier.

Restarting Jupyter Notebook seems to fix it temporarily but soon enough it returns to an error.

After trying basemap for too long with unsolvable errors appearing without any apparent reason, I decide to give up on basemap and find something else.

I switch to cartopy.

26/03/2021 (Week 7) – The Switch to MATLAB

Cartopy has been a complete mess, problems right from installation to the documentation being hard to follow through and a lack of tutorials suitable for this project.

Decide to try geopandas as a last resort but cannot get it installed due to unknown errors.

Eventually decide to give up on jupyter and python altogether due to the amount of errors and time being consumed trying to fix them.

After discussion with my supervisor, I decide to give MATLAB a shot.

02/04/2021 (Week 8) - MATLAB

Taking a long time to install MATLAB on my computer, seems to be a very large download.

Finally installed MATLAB, have to refamiliarize myself with this as it has been a long while since I last coded using MATLAB.

Watching tutorials and trying to code myself.

After a while I seem to have gotten the hang of it enough to try and resume with my project in MATLAB.

09/04/2021 (Week 9) – Preparing the Data

Must now be able to open the data files and extract the needed info in MATLAB.

MATLAB thankfully seems to have built in functions for these, no need to install external libraries to deal with .nc files.

After a while, I manage to open all the files, and am able to extract and store pieces of data as variables in MATLAB.

I realise I need to make some changes to the data to make it usable. Some of the data is in km, other coordinates are in ecef format.

Rakib Khan
Rk704
Supervisor: Biagio Forte

I spend time searching and manually develop functions to make appropriate conversions.

16/04/2021 (Week 10) – Plotting in MATLAB

Need to start visualising the data by making plots, watching tutorials and reading documentation and plotting.

Got some basic plots with the 'plot' function up and running.

Trying to make plots with the actual data is tricky due to x and y variables not being the same size.

Ended up manually cutting some data to make the x and y variables the same size, seems to be working for now.

23/04/2021 (Week 11) – More Plotting and Extra Work

After a discussion with my supervisor, I realised my manually coded functions to convert from ecef to lla were not accurate enough and that natively supported functions exist within MATLAB.

Spend time learning about these functions watching tutorial and reading documentation.

Rewrote unit conversion scripts to now include MATLAB's ecef2lla functions.

Tried out combining various plots of data from different LEOs.

Also plotted delta TEC using the diff function.

30/04/2021 (Week 12) – Forecasting

Need to implement forecasting using machine learning on the data.

Watching tutorials on implement MATLAB's machine learning algorithms.

Rakib Khan
Rk704
Supervisor: Biagio Forte

Deferred Assessment Work

Did more background research

Got book 'Machine Learning Techniques for Space Weather'

Reading though book in more detail

Learning in more detail workings of space weather and its impacts

Learning more about past works by other researchers and authors

Did more data visualisations and implemented 3d plots

Learned up more theory for machine learning in more detail for supervised, unsupervised and reinforcement learning

Learned more about use cases of neural networks for space weather prediction for cases such as solar fares and relativistic electron prediction

Learned and familiarised with regression for machine learning

Implemented regression to data in MATLAB to create best fit line